

IMPROVE MULTI-BASELINE INSAR PARAMETER RETRIEVAL BY SEMANTIC INFORMATION FROM OPTICAL IMAGES

Jian Kang¹, Yuanyuan Wang¹, Marco Körner², Xiao Xiang Zhu^{1,3}

(1) Signal Processing in Earth Observation, Technical University of Munich

(2) Chair of Remote Sensing Technology, Technical University of Munich

(3) Remote Sensing Technology Institute, Germany Aerospace Center

Abstract—One of the most unique benefits of multi-baseline synthetic aperture radar interferometry (InSAR) is the long-term monitoring of subtle ground deformation over large areas. Most state-of-the-art algorithms for retrieving such parameter are based on single pixels, e.g. Permanent Scatterer InSAR [1] or clusters of ergodic pixels with stationary phases e.g. SqueeSAR [2]. None of the studies has addressed the joint inversion in an object level, where the true interferometric phase may be varying subject to topography and deformation. Recently, one study has investigated SAR and optical data fusion in order to make use of the rich semantic information from optical images [3]. Based on that work, we seek to investigate the possibility of an object-level multi-baseline InSAR deformation reconstruction given the semantic information from the corresponding optical images. In this paper, we introduced the tensor model for the multi-baseline InSAR inversion and proposed a maximum *a posteriori* estimator of the deformation parameters by including a spatial prior function in the objective function. Substantial improvement in the deformation estimation is observed in the experiments using both simulated and the real SAR data.

I. INTRODUCTION

Long-term monitoring of millimeter-level deformation over large areas is so far only achievable through multi-baseline InSAR techniques such as Persistent Scatterer Interferometry (PSI) [1], [4], [5] and Differential SAR tomography (D-TomoSAR) [6]–[8]. Through modeling the interferometric phase of the scatterers in SAR images, one is able to reconstruct the 3-D position and the deformation history of each scatterer.

Past research on the multi-baseline InSAR mainly focused on the optimal retrieval of the phase history parameters of individual pixels. Although Distributed Scatterer (DS)-based methods [2], [9] did employ information of multiple pixels, they exploited purely statistical information, assuming the stationarity of the selected samples. Few study addresses the parameter retrieval using explicit geometric information from the image. Only until recently, [10] has demonstrated that by introducing building footprints from OpenStreetMap as a prior knowledge of pixels sharing similar heights, along with joint sparse reconstruction techniques, a highly accurate tomographic reconstruction can be achieved using six interferograms only, instead of the typically required 20–100. Inspired by this, in [11], [12], we investigated the possibility to utilize semantic information obtained from

optical images to reconstruct the deformation parameters in an object level.

This paper is a demonstration of the object-based parameter retrieval algorithm. We give an introduction to the multi-baseline InSAR inversion problem in the form of a tensor, with an additional spatial penalty function. The spatial support of the object is given by the semantic information from optical images. Such semantic classification label is often difficult to obtain from SAR images, because of the complex scattering mechanism. The proposed algorithm is applied to the deformation retrieval of one target bridge area in Berlin.

II. TENSOR-BASED INSAR DEFORMATION RETRIEVAL

A. Tensor model for the multi-baseline InSAR inversion

Given the spatial area of pixels belonging to the same object, the interferometric phase stack can be represented by a 3-mode tensor $\mathcal{G} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$, where I_1 , I_2 represent the spatial dimension in range and azimuth, and I_3 denotes the number of SAR images. Similar to the vector model of Permanent Scatter (PS) time series, the corresponding tensor model is

$$\bar{\mathcal{G}}(\mathbf{S}, \mathbf{P}) = \exp\left\{-j\left(\frac{4\pi}{\lambda r} \mathbf{S} \otimes \mathbf{b} + \frac{4\pi}{\lambda} \mathbf{P} \otimes \boldsymbol{\tau}\right)\right\}, \quad (1)$$

where $\bar{\mathcal{G}}(\mathbf{S}, \mathbf{P})$ is the modeled complex phase tensor of the observed object, \mathbf{S}, \mathbf{P} represent the matrices of the unknown elevation and deformation parameters, respectively, \mathbf{b} is the vector of the spatial baseline, $\boldsymbol{\tau}$ is the vector of the deformation model, e.g. $\boldsymbol{\tau} = \mathbf{t}$ for linear motion, and $\boldsymbol{\tau} = \sin(2\pi(\mathbf{t} - t_0))$ for the seasonal motion model with the temporal baseline \mathbf{t} , λ is the wavelength of the radar transmitted signals and r denotes the range between radar and the observed object. The symbol \otimes is the outer product for the dimension expansion.

Based on this, the Maximum Likelihood Estimator (MLE) of the elevation and deformation parameter matrices assuming Gaussian noise can be expressed in the following tensor form

$$\{\hat{\mathbf{S}}, \hat{\mathbf{P}}\} = \underset{\mathbf{S}, \mathbf{P}}{\operatorname{argmin}} \frac{1}{2} \|\mathcal{G} - \bar{\mathcal{G}}(\mathbf{S}, \mathbf{P})\|_F^2, \quad (2)$$

where \mathcal{G} is the observed InSAR phase tensor, and $\|\cdot\|_F$ is the tensor Frobenius norm.

B. Maximum a posteriori (MAP) estimator in tensor form

Equation (2) extends the MLE from a vector level [13], [14] to a tensor level, which is theoretically identical to a pixel-wise inversion. Assuming certain prior knowledge, e.g. smoothness, sparsity, of the deformation map in the object area, the tensor-based ML estimator can be extended accordingly to the Maximum *a posteriori* (MAP) estimator with the form of:

$$\{\hat{\mathbf{S}}, \hat{\mathbf{P}}\} = \underset{\mathbf{S}, \mathbf{P}}{\operatorname{argmin}} \frac{1}{2} \|\mathcal{G} - \bar{\mathcal{G}}(\mathbf{S}, \mathbf{P})\|_F^2 + \gamma f(\mathbf{S}, \mathbf{P}), \quad (3)$$

where the first term denotes the data fidelity term, which is the Frobenius norm of the log-likelihood between the observed tensor \mathcal{G} and the noiseless tensor model $\bar{\mathcal{G}}$, $f(\mathbf{S}, \mathbf{P})$ denotes the penalty term which describes the spatial priors of \mathbf{S} and \mathbf{P} , and γ controls the balance between these two terms. In this paper, we exploit the smoothness prior of \mathbf{P} . The estimator can be expressed as follows:

$$\{\hat{\mathbf{S}}, \hat{\mathbf{P}}\} = \underset{\mathbf{S}, \mathbf{P}}{\operatorname{argmin}} \frac{1}{2} \|\mathcal{G} - \bar{\mathcal{G}}(\mathbf{S}, \mathbf{P})\|_F^2 + \gamma \sum_{i,j} |\mathbf{P}(i+1, j) - \mathbf{P}(i, j)| + |\mathbf{P}(i, j+1) - \mathbf{P}(i, j)|. \quad (4)$$

III. BRIDGE DETECTION IN OPTICAL IMAGES

This paper demonstrates the proposed method via bridge monitoring. However, the proposed algorithm is generally applicable to other objects. As bridge detection is not the focus of this paper, the associated method in optical images is briefly introduced as follows:

- River classification by training a Support Vector Machine (SVM) on the extracted texture features, e.g. Local Binary Pattern (LBP).
- Refinement of river segmentations by Chan-Vese segmentation.
- α -shape approximations of river segments.
- Bridge detection based on the parametric bridge edge representation

IV. EXPERIMENT

A. Simulated data

We simulate a multi-baseline SAR phase stack with the spatial deformation pattern shown in Figure 1 (Left column). The linear deformation rates range from 1 mm/year to 2.5 mm/year. We choose a spatial baseline comparable to that of TerraSAR-X and a temporal baseline with regular spacing from 0 to 5 (years). The number of images is set to 20. Uncorrelated complex circular Gaussian noise is added to the simulated stacks with signal-to-noise ratio (SNR) of 0 dB and 5 dB. The comparison of the deformation maps estimated by the pixel-wise periodogram (PSI) and by the proposed object-based approach is shown in Figure 1. One can find that the result obtained by the pixel-wise periodogram is much noisier than the proposed approach. Without considering the spatial information of the deformation, the reconstruction result by the pixel-wise periodogram cannot even reflect the

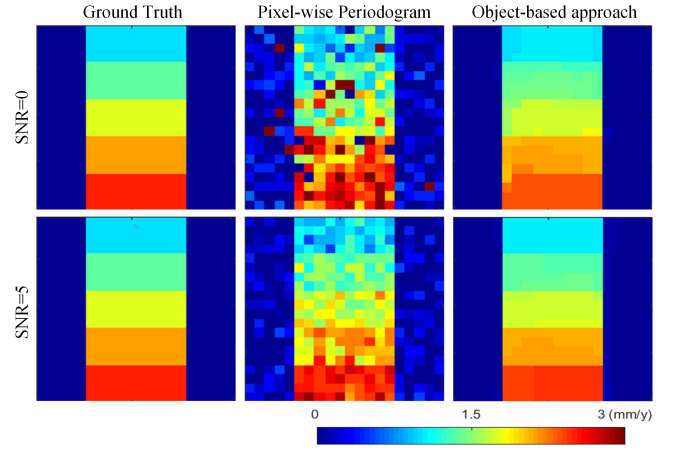


Fig. 1. Comparisons of the linear deformation rates estimated by the pixel-wise periodogram and the proposed object-based approach under two SNRs (0 dB and 5 dB), given the ground truth. (Left) The ground truth deformation map. (Middle) The estimation result by pixel-wise periodogram. (Right) The deformation map reconstructed by the object-based approach.

real deformation pattern under low SNR (0 dB). As shown in Table I, at 0 dB SNR, the proposed approach outperforms the single-pixel periodogram by a factor of 40 in terms of the standard deviation of the estimates.

The parameter γ in equation (3) can be determined by the "L-curve" method, which is a plot of the regularization term with respect to the data fidelity term for a range of regularization parameters [15]. Generally, an L-curve is made up of a flat part and a steep part. The optimal parameter is selected as the corner point of the L-curve where the maximum curvature of the curve is reached. For instance, the L-curve (in log scale) for the simulated InSAR stack with SNR=0dB is plotted in Figure 2. Accordingly, the optimal γ can be determined as 350.

Besides, we have also tested the efficiency of the proposed estimator at different spatial sizes ranging from 20×20 to 40×40 pixels. As illustrated in Figure 3, no significant difference of the efficiency is found, as their SDs are all around 4×10^{-2} (mm/year) under the same noise level (SNR=5dB).

B. Real data

1) *Bridge segmentation in the optical image*: The test area is in central Berlin, which contains several bridges to be monitored. As demonstrated in Figure 4, the size of our study area is 980×1180 with a pixel spacing of 0.2 m. Based on the method introduced in Section III, we classify the associated rivers and bridges, which are marked in blue and green, respectively in Figure 4. As we can see from the result, some building shadows close to the river are misclassified as rivers, since they share similar RGB values with those of rivers. Some bridges do show irregular shapes, especially the top one, since the bridge mask depends on the boundary of the river segments. Yet, this does not affect much the bridge monitoring, since the bridge masks cover most part of the bridges.

TABLE I
NUMERICAL PERFORMANCE FOR THE RESULTS SHOWN IN FIGURE 1

		pixel-wise peridogram	the proposed approach
Standard Deviation (SD) [mm/year]	SNR = 0 dB	2.64	6.32×10^{-2}
	SNR = 5 dB	2.31×10^{-1}	3.94×10^{-2}

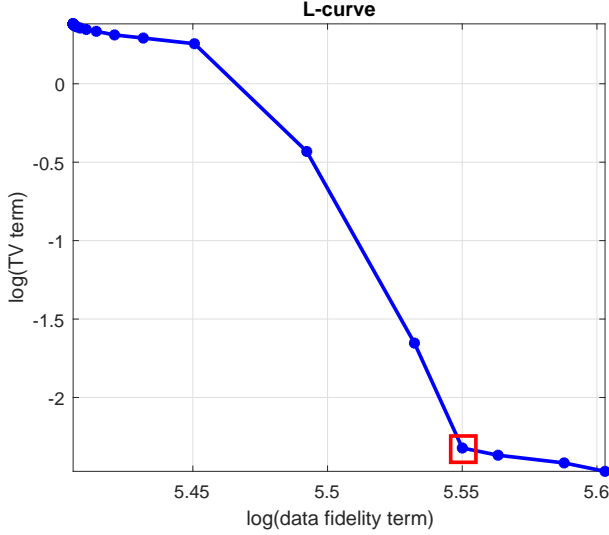


Fig. 2. The L-curve plot (in log scale) of the regularization term with respect to the data fidelity term in equation (4) for a range of γ . The optimal γ can be automatically chosen by finding the maximum curvature of the curve. In this simulation (SNR=0dB), γ is accordingly selected as 350 (red rectangle).



Fig. 3. The plot of the errors of our result w.r.t. different spatial sizes ranging from 20×20 to 40×40 pixels. It is shown that there is no significant difference in the efficiency, as their SDs are all around 4×10^{-2} (mm/year) under the same noise level (SNR=5dB).

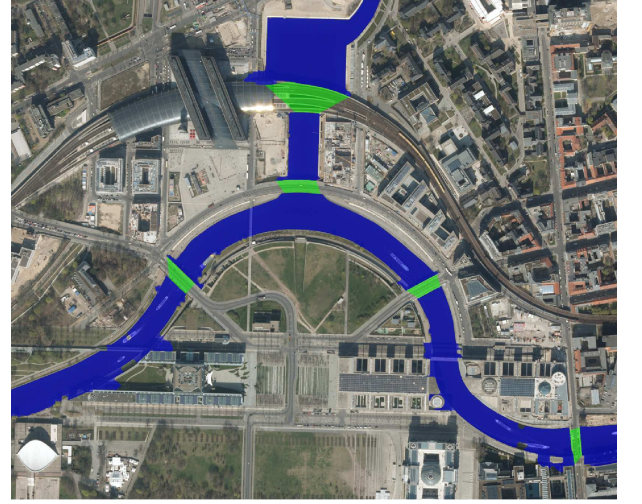


Fig. 4. The classifications of the rivers and bridges by the proposed approach, which are masked in blue and green, respectively. As we can see from the result, some building shadows are misclassified as rivers, since they share similar RGB values with those of rivers. Some bridges do show irregular shapes, especially the top one, since the bridge mask depends on the boundary of the river segments. Yet, this does not affect much the bridge monitoring, since the bridge masks cover most part of the bridges.

2) *Bridge mask projected to the SAR image:* Given the classified bridge areas in optical images, we can obtain the corresponding bridge masks in the SAR image based on the 'SARptical' work in [3]. For example, the top bridge mask in Figure 4 projected in the SAR image is shown in Figure 5. The area extracted for the deformation reconstruction is chosen by the bounding box of the bridge mask as indicated by the red rectangle. The dimension of this dataset is 243×66 pixels with 109 SAR images in total.

3) *Object-based InSAR deformation retrieval:* The selected bridge undergoes a seasonal motion which is primarily caused by the thermal dilation of the steel railways on the bridge. The estimated amplitudes of this periodic motion using the proposed method and the pixel-wise periodogram are shown in Figure 6. Consistent with the analysis of the simulation, more incorrectly estimated deformation points are shown in the pixel-wise periodogram result. On the contrary, the object-based approach can obtain much more reliable deformation reconstruction of the study area.

V. CONCLUSIONS

In this paper, we proposed a general framework for object-based multi-pass InSAR parameters estimation, i.e. introducing a spatial regularization term based on given object labels. To demonstrate the application of the proposed

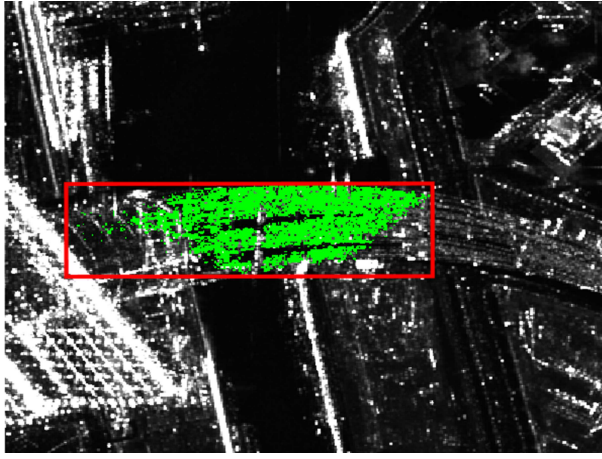


Fig. 5. The bridge mask (green) in the SAR image obtained by projecting the corresponding mask from the optical image [3], which is the top bridge shown in Figure 4, and the red rectangular area is used for the object-based deformation reconstruction.

method for bridge monitoring, we also proposed a bridge detection method in optical images.

The experiments show that the regularization parameter γ does not sensitively influence the efficiency of the reconstruction result, especially at the high SNR (5dB). Its optimal value can be automatically determined by the "L-curve" method within a given range of γ . This range setting depends on the dataset. We find it to be 100 – 350 in our simulated experiments. In addition, we observe that the efficiency of the proposed method does not show significant change with respect to different object dimensions, which indicates that the spatial size can be kept small for large area parallel processing.

To summarize, the proposed robust object-based approach is a novel framework that combines geometric information and multi-pass InSAR technique. It is suited for areas with homogenous pixels as well as for urban areas where the pixels are highly nonergodic. The proposed approach can be efficiently solved by smart optimization methods which renders it suitable for operational processing.

REFERENCES

- [1] A. Ferretti, C. Prati, and F. Rocca, "Permanent scatterers in SAR interferometry," *IEEE Transactions on geoscience and remote sensing*, vol. 39, no. 1, pp. 8–20, 2001.
- [2] A. Ferretti, A. Fumagalli, F. Novali, C. Prati, F. Rocca, and A. Rucci, "A new algorithm for processing interferometric data-stacks: SqueeSAR," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 9, pp. 3460–3470, 2011.
- [3] Y. Wang, X. X. Zhu, B. Zeisl, and M. Pollefeys, "Fusing meter-resolution 4-D InSAR point clouds and optical images for semantic urban infrastructure monitoring," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 1, pp. 14–26, 2017.
- [4] S. Gernhardt and R. Bamler, "Deformation monitoring of single buildings using meter-resolution SAR data in PSI," *ISPRS journal of photogrammetry and remote sensing*, vol. 73, pp. 68–79, 2012.
- [5] M. K. Bert, "Radar Interferometry: Persistent Scatterers Technique," *The Netherlands: Springer*, 2006.
- [6] X. X. Zhu and R. Bamler, "Very high resolution spaceborne SAR tomography in urban environment," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 12, pp. 4296–4308, 2010.

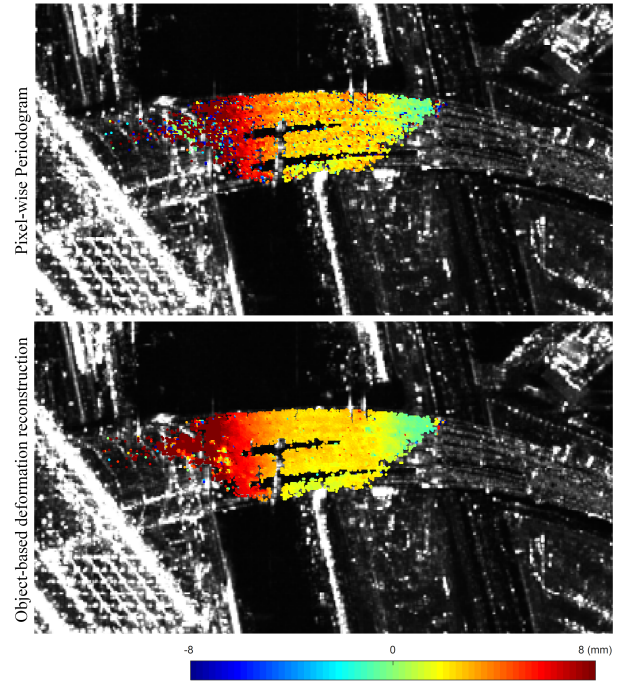


Fig. 6. The amplitudes of seasonal motion estimated by the single-pixel based periodogram and the proposed object-based approach of the same bridge area. Consistent with the analysis of the simulation, more incorrectly estimated deformation points are shown in the pixel-wise periodogram result. On the contrary, the object-based approach can obtain much more reliable deformation reconstruction of the study area.

- [7] G. Fornaro, S. Verde, D. Reale, and A. Pauciuillo, "CAESAR: An approach based on covariance matrix decomposition to improve multibaseline-multitemporal interferometric SAR processing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 4, pp. 2050–2065, 2015.
- [8] X. X. Zhu and R. Bamler, "Tomographic SAR inversion by L1 norm regularization—The compressive sensing approach," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 10, pp. 3839–3846, 2010.
- [9] Y. Wang, X. X. Zhu, and R. Bamler, "Retrieval of phase history parameters from distributed scatterers in urban areas using very high resolution SAR data," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 73, pp. 89 – 99, 2012.
- [10] X. X. Zhu, N. Ge, and M. Shahzad, "Joint Sparsity in SAR Tomography for Urban Mapping," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 8, pp. 1498–1509, Dec 2015.
- [11] J. Kang, Y. Wang, M. Körner, and X. X. Zhu, "Object-based InSAR deformation reconstruction with application to bridge monitoring," in *Geoscience and Remote Sensing Symposium (IGARSS), 2016 IEEE International*. IEEE, 2016, pp. 6871–6874.
- [12] J. Kang, Y. Wang, M. Krner, and X. X. Zhu, "Robust Object-Based Multipass InSAR Deformation Reconstruction," *IEEE Transactions on Geoscience and Remote Sensing*, vol. PP, no. 99, pp. 1–13, 2017.
- [13] A. De Maio, G. Fornaro, and A. Pauciuillo, "Detection of single scatterers in multidimensional SAR imaging," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 7, pp. 2284–2297, 2009.
- [14] Y. Wang and X. X. Zhu, "Robust estimators for multipass SAR interferometry," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 2, pp. 968–980, 2016.
- [15] P. C. Hansen and D. P. O'Leary, "The use of the L-curve in the regularization of discrete ill-posed problems," *SIAM Journal on Scientific Computing*, vol. 14, no. 6, pp. 1487–1503, 1993.